Public Health England

Protecting and improving the nation's health

National Cancer Intelligence Network

Cancer Outcomes and Services Dataset (COSD) Version 6.0

Specification

About Public Health England

Public Health England exists to protect and improve the nation's health and wellbeing, and reduce health inequalities. It does this through world-class science, knowledge and intelligence, advocacy, partnerships and the delivery of specialist public health services. PHE is an operationally autonomous executive agency of the Department of Health.

Public Health England
Wellington House
133-155 Waterloo Road
London SE1 8UG

Tel: 020 7654 8000
www.gov.uk/phe
Twitter: @PHE_uk
Facebook: www.facebook.com/PublicHealthEngland

Any enquiries regarding this publication should be sent to us at publications@phe.gov.uk

Published February 2015
PHE publications gateway number: 2014605

The intelligence networks

Public Health England operates a number of intelligence networks, which work with partners to develop world-class population health intelligence to help improve local, national and international public health systems.

National Cancer Intelligence Network

The National Cancer Intelligence Network (NCIN) is a UK-wide initiative, working to drive improvements in cancer awareness, prevention, diagnosis and clinical outcomes by improving and using the information collected about cancer patients for analysis, publication and research.

National Cardiovascular Intelligence Network

The National cardiovascular intelligence network (NCVIN) analyses information and data and turns it into meaningful timely health intelligence for commissioners, policy makers, clinicians and health professionals to improve services and outcomes.

National Child and Maternal Health Intelligence Network

The National Child and Maternal Health Intelligence Networks (NCMHIN) provides information and intelligence to improve decision-making for high quality, cost effective services. Their work supports policy makers, commissioners, managers, regulators, and other health stakeholders working on children's, young people's and maternal health.

National Mental Health Intelligence Network

The National Mental Health Intelligence Network (NMHIN) is a single shared network in partnership with key stakeholder organisations. The Network seeks to put information and intelligence into the hands of decision makers to improve mental health and wellbeing.

National End of Life Care Intelligence Network

The National End of Life Care Intelligence Network (NEoLCIN) aims to improve the collection and analysis of information related to the quality, volume and costs of care provided by the NHS, social services and the third sector to adults approaching the end of life. This intelligence will help drive improvements in the quality and productivity of services.

Title	Cancer Outcomes and Services Dataset v6.0 Specification		
SCCI reference	SCCI 1521 Amd 17/2014		
Sponsor	Jane Allberry		
Senior Responsible Officer	Sean Duffy	Status	Final
Developer	Trish Stokes	Version	6.2
Author	Trish Stokes/Carlos Rocha	Version Date	30/01/2015

Amendment history:

Version	Date	Amendment history
6.0	18.12.2014	Final for submission for 2015 changes(Versioning aligned with Dataset and schema versions)
6.1	16.01.2015	Added Appendix I
6.2	30.01.2015	Implementation completion date amended to 1st January 2016 Minor amendments following ISAS appraisal meeting Correction of item deleted in error

Approvals:

This document MAY be approved by the following:

Name	Signature	Title / Responsibility	Date	Version
COSD Governance Board	COSD Governance Board	COSD Governance Board	16.02.2015	6.2

Glossary of terms:

Term	Acronym	Definition
Burden Advice and Assessment Service	BAAS	The Burden Advice and Assessment Service (BAAS) has taken over many of the functions of the Review of Central Returns (ROCR) programme regarding the burden of data collections. This includes advice, assessment and review of new and existing collections and standards as well as recommendations to the Standardisation Committee for Care Information (SCCI).
Cancer		For the purposes of this standard the term 'cancer' is used throughout the standard and related documents to cover all conditions defined as registerable by the UK and Ireland Association of Cancer Registries.
Cancer Centres		Organisations which help people to live with, through and beyond cancer by bringing together specialist clinical and professional staff and communities of support.
Cancer Outcomes and Services Dataset	COSD	The COSD is the national standard for reporting cancer in the NHS in England. It replaced the previous National Cancer Dataset and includes the former Cancer Registration dataset and additional site specific data items relevant to the different tumour types.
Cancer Registration Dataset	CRDS	The dataset requirements for cancer registration now incorporated into the COSD.
Cancer Registries		Organisations which exist internationally to collect, process, analyse and disseminate data on cancer patients in their local regions.
Care Quality Commission	CQC	One of the independent regulators of health & social care in England.
Commissioners		Organisations that plan, purchase and monitor services to meet the health needs of their local population.

Term	Acronym	Definition
Diagnostic Imaging Dataset	DIDS	Dataset containing diagnostic imaging test activity across the NHS, taken from Radiology Information Systems.(See ISB Standard 1577)
Extensible Markup Language	XML	Extensible Markup Language (XML) is a set of rules for encoding documents in machine-readable form.
Health and Social Care Information Centre	HSCIC	The Health and Social Care Information Centre is England's central, authoritative source of health and social care information for frontline decision makers, which builds upon the Health and Social Care Act 2012.
Improving Outcomes: A Strategy for Cancer	IOSC	The overarching strategy for cancer services in England.
Information Standard	IS	A measure that ensures that information is managed in a consistent manner across health and social care, both by the computers and the staff.
International Statistical Classification of Diseases and Related Health Problems	ICD	A medical classification list for the coding of diseases, signs and symptoms, abnormal findings, complaints, social circumstances and external causes of injury or diseases, as maintained by the World Health Organization (WHO). (The title is followed by the revision number, eg ICD10 is the tenth revision.)
International Classification of Diseases for Oncology	ICD-O	An extension of the ICD coding system used principally in tumour or cancer registries for coding the site (topography) and the histology (morphology) of neoplasms. (The title is followed by the revision number, eg ICD-O-3 is the third revision.)
National Cancer Dataset	NCDS	The previous nationally approved reference standard for the collection of cancer data now incorporated into the COSD.

Term	Acronym	Definition
National Cancer Intelligence Network	NCIN	A UK-wide initiative, working to drive improvements in standards of cancer care and clinical outcomes by improving and using the information collected about cancer patients for analysis, publication and research. NCIN is one of a number of Health Intelligence Networks operated by Public Health England.
National Cancer Registration Service	NCRS	The NCRS is the national cancer registration service for England collecting cancer data from all NHS Providers of cancer care in England. The NCRS is a function within the National Disease Registration Service within Public Health England.
National Cancer Waiting Times Monitoring Dataset	NCWTMDS	The Information Standard (ISB0147) used to monitor the time that patients with suspected and diagnosed cancer have wait for appointments, tests and treatments.
Providers		Organisations that provide health services.
Public Health England	PHE	Public Health England is an executive agency of the Department of Health in the United Kingdom, taking up its full powers from 1 April 2013 Its role is protecting and improving the nation's health and well being and to reduce inequalities.
Radiotherapy Dataset	RTDS	A standard dataset covering every patient treated with radiotherapy in the NHS in the England.
Royal College of Pathologists	RC Path	A professional membership organisation committed to setting and maintaining professional standards and to promoting excellence in the practice of pathology.
Systemic Anti Cancer Therapy Dataset	SACT	The national collection of all cancer chemotherapy data in the NHS in England, which covers all solid tumour and hematological malignancies. This includes all adult and paediatric cancer patients, those in clinical trials, and covers acute inpatient, day case, outpatient and community settings.

Term	Acronym	Definition
United Kingdom and Ireland Association of Cancer Registries	UKIACR	The UKIACR brings together organisations with an interest in developing cancer registration as a resource for studying and controlling cancer in the UK and Ireland.
XML schema		The documentation, definitions and descriptions required to enable the production and transmission of data for a specific XML language.

Contents

1. Overview

1.1 Summary

The table below contains a summary of the information standard.

Standard	
Standard Number	SCCI1521
Standard Title	Cancer Outcomes and Services Dataset
Description	The Cancer Outcomes and Services Dataset (COSD) is a compiled dataset which provides the standard for secondary uses information required to support implementation and monitoring of Improving Outcomes: a Strategy for Cancer (IOSC). This standard consists of: A set of individual data items, with their definitions. The assemblage of these data items into collection areas. The means of flowing the data items. Compilation of the data items into a single reconciled and verified dataset. The COSD has replaced the former National Cancer Dataset (NCDS) (including the approved site specific data Items such as those for breast, colorectal, lung, head and neck, urological, upper gastrointestinal, gynaecological, sarcoma and skin cancers), and the former Cancer Registration Dataset. It incorporates the National Cancer Waiting Times Monitoring Dataset (NCWTMDS) and items from the Systemic Anti Cancer Therapy Dataset (SACT) and the Radiotherapy Dataset (RTDS), all of which also remain as separate standards. All patients diagnosed with or receiving cancer treatment in or funded by the NHS in England are covered by the standard. This includes adult and paediatric cancer patients. Providers of cancer services are required to provide a monthly return on all cancer patients diagnosed from 1 January 2013 using this dataset. Data is collated via the NCRS local offices, and former

Standard	
	mechanisms for transmission of data from Providers to registries has been extended to carry the COSD dataset.
Applies to	Cancer centres, cancer units and all other providers of NHS commissioned cancer services Developers and suppliers of electronic systems for use in NHS commissioned cancer centres and NHS provider services Organisations purchasing electronic systems for use in NHS commissioned cancer centres and NHS provider services Users of secondary data about cancer at both national and local levels, including: At a national level: the Department of Health (DH), National Cancer Registration Service (NCRS), National Cancer Intelligence Network (NCIN) and other appropriate national information, research and service planning organisations, eg The Health and Social Care Information Centre (HSCIC), Care Quality Commission (CQC), Monitor, Public Health England (PHE), NHS Improving Quality (NHS IQ) At a local level: Strategic Clinical Networks (SCNs) and local Cancer Service Provider Networks, the local NCRS offices, commissioners and provider organisations will have data on cancer services based on this national standard.
Release	
Release Number	Amd 17/2014
Release Title	Version 6.0
Description	This is a change to the standard which introduces some amendments to the current dataset, an extension of scope and a revision of the current schema specification in order to continue to meet the business objectives of the standard. There has also been an extension to the deadline for submitting data in XML structured format agreed for pathology in recognition of the difficulties encountered in meeting the initial deadline of January 2015. The deadline has now been extended to 1 January 2016 and providers are asked to provide a plan to the NCRS for meeting this requirement by July 2015.
Implementation start and completion date	Between 1 April 2015 and 1 January 2016.

1.2 Supporting documents

This specification should be read in conjunction with the following documents:

Product	Document Reference	Title
Change Request	www.hscic.gov.uk/isce/publication/scci1521	Cancer Outcomes and Services Dataset (COSD) V6.0 - Change Request
Dataset	http://www.ncin.org.uk/collecting_and_using_data/data_collection/cancer_outcomes_and_services_dataset_cosd_latest_downloads	COSD Dataset v6.0
Implementation User Guide	TBA	COSD User Guide v6.0
DD Change Paper	TBA	cr-1492-2014-12-01

1.3 Related standards

This specification should be read in conjunction with the following standards:

Ref #	Reference	Title
ISB 0147	http://www.isb.nhs.uk/documents/isb-0147/amd-23-2011/index_html	Cancer Waiting Times Monitoring Dataset
ISB 0111	http://www.isb.nhs.uk/documents/isb-0111/dscn-22-2008/index_html	Radiotherapy Dataset
ISB 1533	http://www.isb.nhs.uk/documents/isb-1533/amd-63-2010/index_html	Systemic Anti Cancer Therapy Dataset
ISB 1577	http://www.isb.nhs.uk/documents/isb-1577/amd-10-2011/index_html	Diagnostic Imaging Dataset
ISB 0021	http://www.isb.nhs.uk/documents/isb-0021/amd-86-2010/index_html	International Classification of

Ref #	Reference	Title
		Diseases
ISB 0034	http://www.isb.nhs.uk/documents/isb-0034/amd-26-2006/index_html	SNOMED CT
n/a	http://www.rcpath.org/publications-media/publications/datasets/datasets-TP.htm	Royal College of Pathologists Standards and Datasets for Histopathology Reporting on Cancers and Tissue Pathways

1.4 Contacts

Sponsor	
Name	Jane Allberry
Organisation	Department of Health
Senior responsible officer	
Name	Sean Duffy
Organisation	NHS England
Developer	
Name	NCIN
Organisation	National Cancer Intelligence Network
Email Address	cosd@ncin.org.uk
Implementation manager	
Name	Stephen Raynor
Organisation	National Cancer Registration Service, Public Health England
Email Address	Stephen.Raynor@phe.gov.uk
Maintenance manager	
Name	Trish Stokes
Organisation	National Cancer Intelligence Network, Public Health England
Email Address	trish.stokes@phe.gov.uk

1.5. Status of Documents

Assurance that this information standard is appropriate for the use specified in the specification document has been provided by the Standardisation Committee for Care Information (SCCI), a sub-group of the National Information Board.

This information standard comprises the following documents:

- specification
- implementation guide

In addition, a change request provides an overview of the change from the last version of the standard.

An Information Standards Notice (SCCI1521 Amd 17/2014) has been issued as a notification of use and implementation timescales. Please read this alongside the listed documents.

The controlled versions of these documents can be found on the HSCIC website.

2. Health and care organisations

2.1 Requirements

#	Requirement
1	NHS Providers of cancer services (hereinafter referred to as NHS providers) **should** read the specification in conjunction with the implementation guidance to identify how the standard is applicable to them.
2	NHS providers **must** review their system compatibility against this standard to identify any changes required to current practice to ensure that all data items in COSD v6.0 can be flowed electronically by dates specified in the implementation guidance.
3	NHS providers **must** submit the data using the XML format by January 2015 for extracts from MDT management systems.
4	NHS providers **must** submit the data using the XML format by 1 January 2016 for extracts from pathology systems.
5	NHS providers **must** provide an action plan to the NCRS by July 2015 for how they will meet the target to submit the data using the XML format by 1 January 2016 for extracts from pathology systems.
6	While developing the XML format for data submission NHSpProviders of cancer services **may** agree interim arrangements with the local NCRS office to submit data in alternative electronic formats providing the NCRS is able to extract the specified data from these.
7	NHS Providers **should not** utilise this dataset primarily to support their clinical and operational data capture.

2.2 Conformance criteria

Organisational type	#	Criteria
Providers	1	NHS Providers MUST submit COSD v6.0 data items as specified in the Implementation Guidance within the defined time period and in the format specified in these documents.
	2	All submitted data files MUST contain the specified linkage items at record level to enable linkage of the relevant cancer registration records.
	3	NHS Providers MUST submit the agreed data items within 25 working days of the month end following diagnosis date.

Organisational type	#	Criteria
	4	NHS providers **must** submit the agreed data items within 25 working days of the month end following treatment start date.
	5	NHS providers **must** submit further records for all cases within 25 working days of the month end following any additional or amendments to the data items.
	6	NHS providers **must** agree methods of submission with the NCRS for all items not flowed as part of the standard extract.
	7	NHS providers **must** notify the NCRS of any known reasons for significant variation in the number of new cases submitted monthly if applicable.
	8	NHS providers **must** review monthly feedback from the NCRS.
	9	NHS providers **must** audit case ascertainment, quality and completeness on receipt of quarterly feedback reports from the NCRS and notify the NCRS of reasons for any discrepancies.
	10	Providers **must** report a minimum of 80% of all expected cases annually by site specific tumour group as agreed with the NCRS.
	11	NHS providers **must** submit the dataset extracted from their MDT management systems in XML.
	12	NHS providers **must** submit an action plan to the NCRS by July 2015 outlining how they will achieve submission of data from pathology systems in XML by 1 January 2016

Conformance is measured against the COSD conformance framework which has been published on the COSD webpage. Basic feedback on conformance is provided to providers through the NCRS's COSD conformance reporting portal (through NHS N3 network only), which is under ongoing development to provide more detailed feedback. Feedback reports are also provided to the COSD Governance Board in order to monitor and manage compliance to the information standard.

3. IT systems suppliers

3.1 Requirements

#	Requirement
1	Suppliers of Cancer IT systems MUST implement changes in accordance with their local contractual arrangements to enable all specified data items in COSD v6.0 to be captured and extracted in compliance with the Specification and Implementation Guidance.

3.2 Conformance Criteria

Criteria
The requirement above MUST be met.

4. Scope

4.1 In scope

The dataset relates to all cancer patients, both adult and paediatric, in acute inpatient and outpatient settings. The trigger for data collection is when a diagnosis, or suspected diagnosis of cancer is confirmed, primarily this diagnosis takes place within secondary care.

The standard covers neoplasms coded within ICD-10 diagnosis codes range C00 – C97, D00 – D48 and E85.9[1]. (See Appendix A for list of mandatory registerable conditions (UKIACR Library of Recommendations)).

Following the specific inclusion of information relating to the management of patients with either recurrent or metastatic breast cancer in 'Improving Outcomes: A Strategy for Cancer', a specific subset of data items has been identified and is included in the implementation guidance to support the collection of data for this group of patients.

In addition to the data required for recurrent or metastatic breast cancer patients, the standard has been extended and now includes as a minimum demographic and diagnosis details for all cancer recurrences.

4.2 Out of scope

As a secondary uses dataset, this standard does not define record level data to be used in the delivery of care. The data for COSD should be derived from patient identifiable data which are already recorded for the purpose of care management.

[1] Although Primary amyloidosis (E85.9) is listed as an E ICD code in the World Health Organization (WHO) disease classification, amongst clinicians it is widely acknowledged and subsequently treated as a cancer, receiving Chemotherapy in cases. While we await the WHO disease classification being updated to reflect this fact, we have extended the scope of the COSD to include this. The United Kingdom and Ireland Association of Cancer Registries (UKIACR) is currently considering its inclusion in the UKIACR Library of Recommendations, which we have referenced in Appendix A.

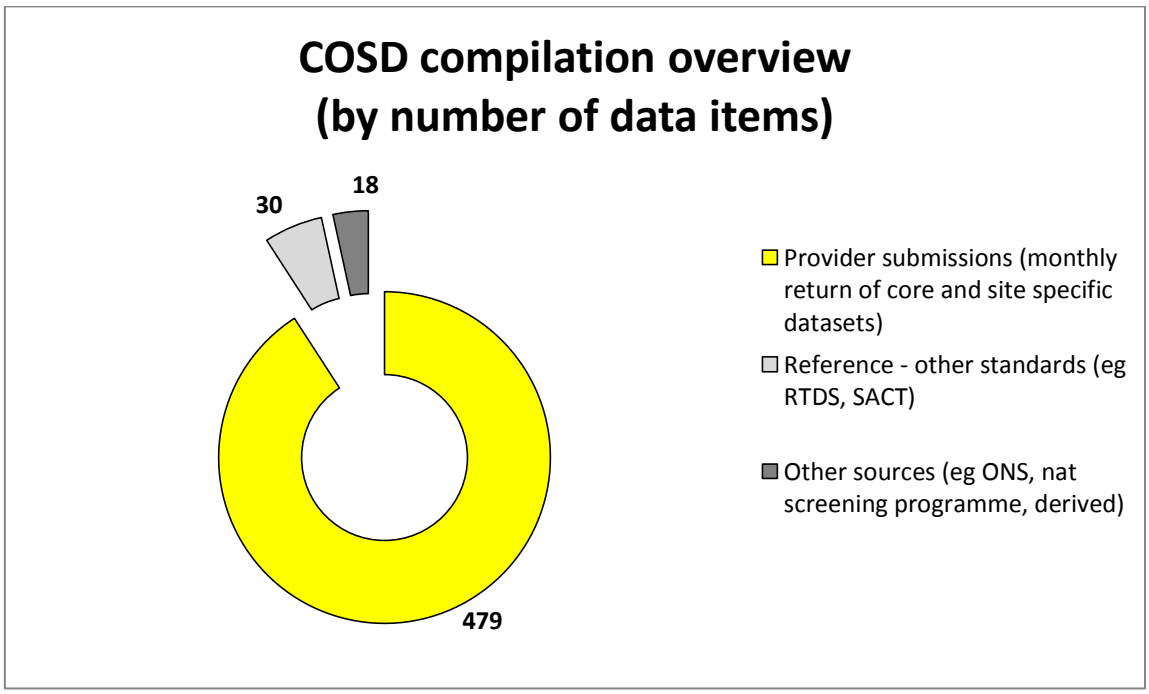
- General Practice – developments in the extraction of data from general practice systems are the subject of other work by the HSCIC, specifically the General Practice Extraction Service (GPES)
- there are no formal links at present but this is expected to be developed in the future
- radiotherapy - this is subject to an existing standard, the National Radiotherapy Dataset (RTDS) - ISB 0111
- chemotherapy – this is subject to an existing standard, the systemic anti-cancer therapy dataset (SACT) - ISB 1533
- imaging – this is subject to an existing standard, the diagnostic imaging dataset (DID) - ISB 1577

5. Implementation and use

5.1 Guidance

5.1.1 High level view

This standard, together with the supporting COSD dataset v6.0, defines the complete set of secondary uses cancer data for reporting and specifies the items which need to be returned directly by NHS providers. Other items are either subject to other standards (such as RTDS), provided from other sources (such as ONS or screening services) or derived (these are clearly distinguished in the dataset documentation). Further details are provided in the user guidance and the following provides summary information only.



Provider submissions: the 479 items shown above comprise the subset of the COSD to which the remainder of this specification refers unless otherwise stated. These are the items which are included in the XML schema and are expected to be flowed directly from NHS providers to the local NCRS branch office, from one or more electronic systems within the provider organisation.

Reference - Other standards: unless otherwise specified, these items are excluded from the remainder of this specification document and are identified separately within the full dataset. These are items which are essential to compile the full dataset but are covered by other information standards and

therefore will not need to be included in the direct data flows for COSD. These items are flowed to the NCRS from other national collection systems.

Other sources: unless otherwise specified, these items are excluded from the remainder of this specification document and are identified separately within the full dataset. These are items which are essential to compile the full dataset but will be collected from other sources and therefore will not need to be included in the direct data flows for COSD.

5.1.2 Model data flow diagram

The following diagram demonstrates how the full COSD dataset will be compiled centrally by local NCRS offices from data flowing from a number of systems and sources. (This diagram includes reference – other standards and other sources data items.)

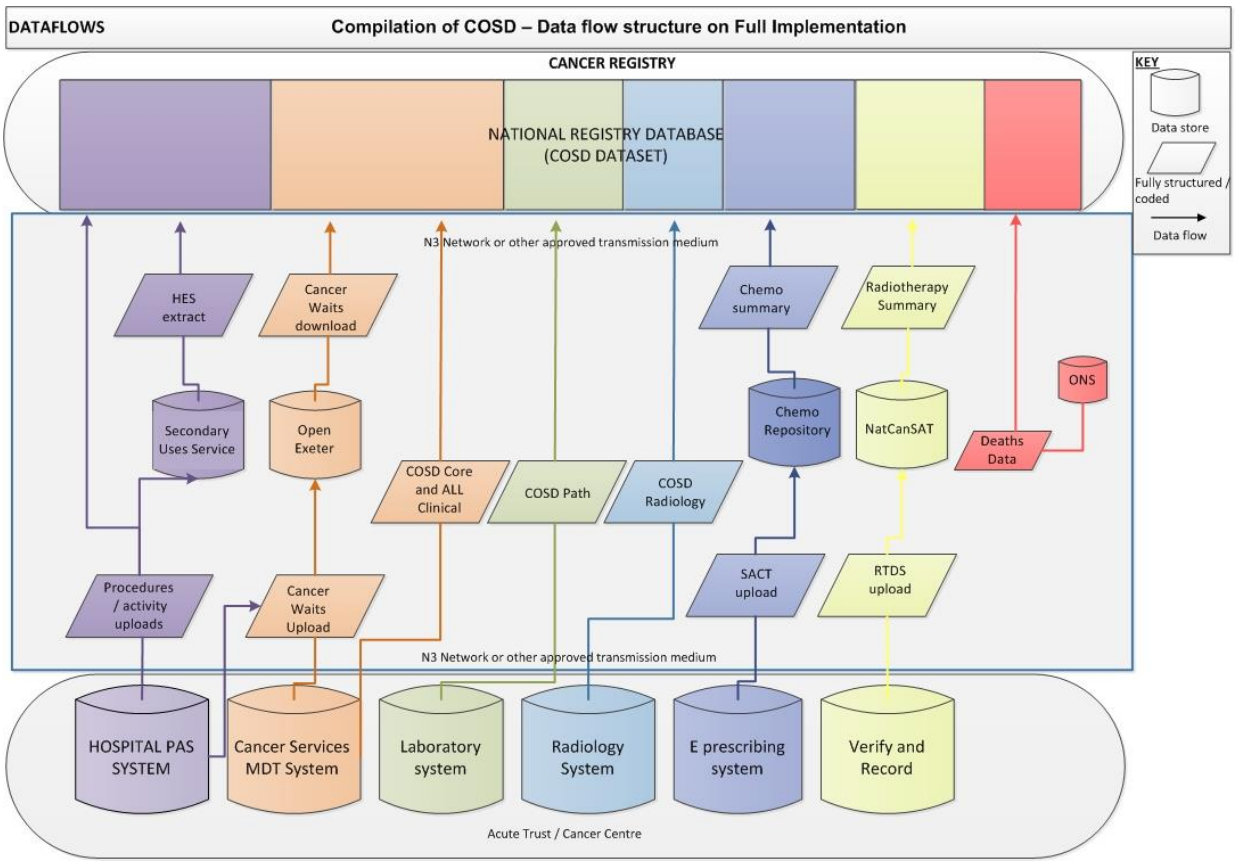


Fig 2: COSD compilation – data flow structure on full implementation

5.1.3 Structure of dataset – provider submissions

There is a core dataset most of which is applicable to all cancers and an additional site specific dataset for each of the 12 identified tumour groups. Some of these site specific datasets contain further subsets applicable to individual diseases. Each recorded case will therefore have a core and usually a site specific dataset completed.

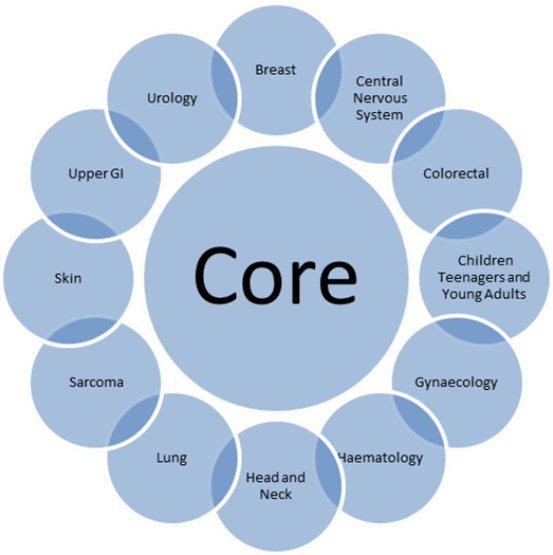


Fig 3: Structure of COSD

5.1.4 Dataset subsections

Within each of the core and site specific datasets, the data items are further grouped according to their stage along the patient pathway as follows:

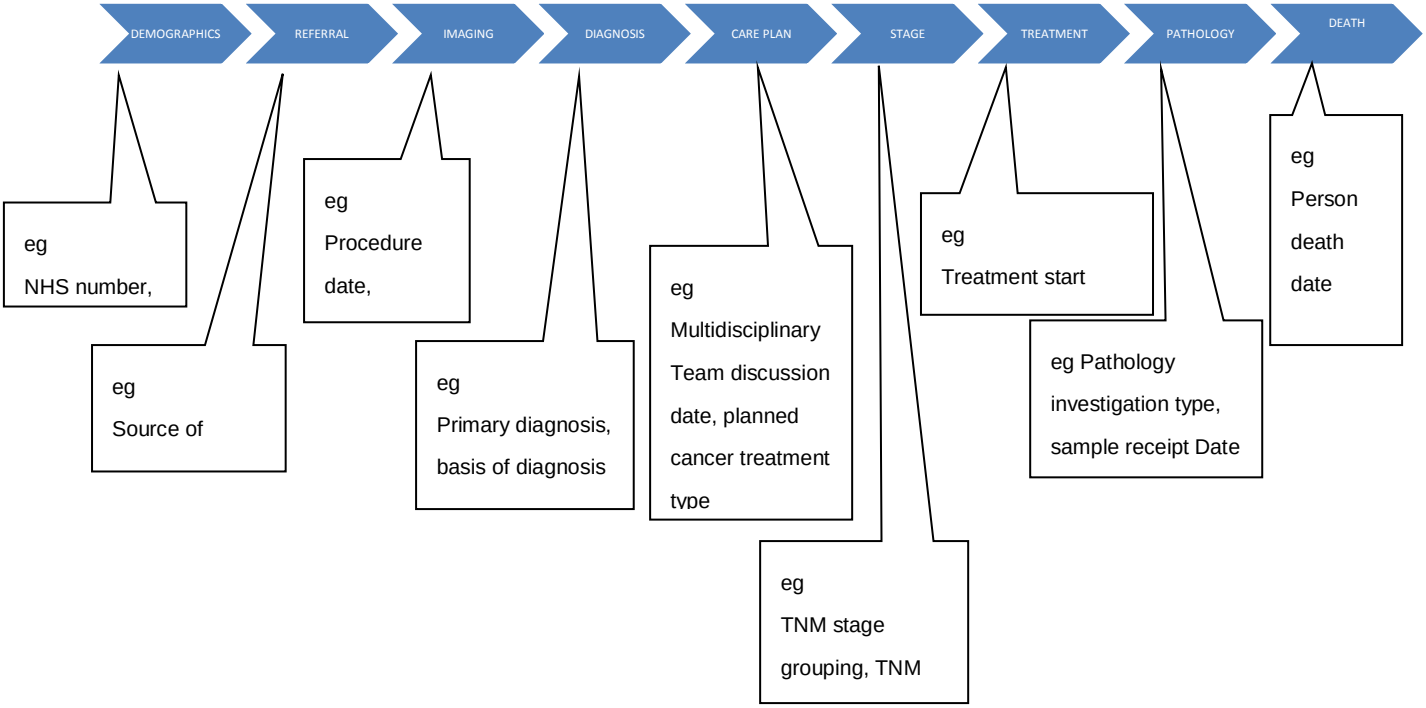


Fig 4: Dataset subsections

5.2 Governance

5.2.1 Information governance

Data collection from all the new sources required to support this expansion of the current cancer registration dataset are covered by existing permissions granted by the Confidentiality Advisory Group (CAG) of the Health Research Authority (HRA).

The dataset contains sensitive and patient-identifiable information items. The CAG of the HRA has confirmed that reporting of patient identifiable data is covered by the NCRS existing support under the Health Service (Control of Patient Information) Regulations 2002 (see Appendix I for details). Reported data will be managed by the NCRS where there is long standing expertise in managing large volumes of confidential data.

Although the data items which are flowed to the NCRS have changed, the data flows (ie which organisations will be receiving the data in identifiable form) remain unchanged. In compliance with the fair processing requirement within the Data Protection Act, provider organisations are expected to inform patients of this purpose for reporting their information and of the potential use of the information for service development, analysis and statistical research.

The UKACR has expanded to include the Republic of Ireland Cancer Registry and consequently has changed its name to the United Kingdom and Ireland Association of Cancer Registries – UKIACR. The policies referred to below will be revised in due course.

The NCRS and Cancer Research UK have developed a patient information leaflet (see Appendix B) that is a useful resource for organisations wanting to develop or revise local information materials. NCRS, as part of PHE, will come under the Department of Health's Data Protection Act registration with the Information Commissioner's Office (ICO). The NCRS is currently reviewing and harmonising its information governance policies to correlate them with those of PHE and maintain compliance with the Health and Social Care Information Centre information governance toolkit. These policies inform for example: access controls of data, server security, encryption and data transfer procedures.

5.2.2 Consent process

Where patients have requested their data are not shared, the provider organisation must ensure that their records are not included in the data downloads submitted to the NCRS It is suggested that a dissent (ie the proactive expression of dissent by an individual from whom consent has not been obtained) or a similar flag is provided in the Provider organisations systems so that the record can then be omitted from the monthly upload.

The NCRS has published an updated patient information leaflet (Appendix B) which explains that individuals have the right to access and have their own data held in the NCRS removed, and which explains the process. If a patient discovers that their information has been uploaded to the NCRS and they wish for this to be deleted, the requester must complete a removal request form (available from the weblink below) and send this to the NCRS (instructions on the form) to action. The NCRS will then, as far as is possible, remove the patient from the NCRS database.

The NCRS information for patients wishing to have their information removed as far as possible from the NCRS database is available on the NCRS website

(http://www.ncr.nhs.uk/removal-request). (see Appendix C for further information).

The NCRS has subject access request policies relating to requests for patients to view their own data. These will be standardised as part of the registry modernisation programme. A sample policy is attached at Appendix F.

5.2.3 Data retention

The NCRS, as part of PHE, is developing a new data retention policy. This will be in place by 1 April 2015. Until this is approved the UKACR policy on data retention and disposal will apply to this data (see Appendix D). For further information please contact the NCRS IG Lead, Roger Hartley: roger.hartley@phe.gov.uk

5.2.4 Data disclosure

The NCRS adheres to the requirements of the Data Protection Act 1998 with regard to the receipt, storage and transfer of information relating to individuals. When releasing data to third parties, all UK registries strictly comply with the Approved UKACR policies on the release of patient-identifiable and potentially identifiable information. Recipients of such data are required to sign a declaration stating that they will protect the information they are entrusted with, use it only for the purpose for which it was supplied and make no attempt to identify information pertaining to particular individuals or to contact individuals. They are also prohibited from presenting any information that may identify an individual. This is also the case with publications produced by NCRS, which present aggregated data only.

The UKACR policy on data disclosure applies to this data and is available (see Appendix E for further information). This continues to apply until replaced by a PHE policy.

5.2.5 Clinical governance

This is a secondary uses standard – no direct patient safety hazards were identified for the dataset itself. Consultation, piloting, user guidance and validation processes address data quality issues that might have an indirect impact on services, patient care and treatment. The risk that patient identifiable data could be accessed or disclosed inappropriately is addressed in the implementation guidance section of the specification and guidance. The risk that the dataset could be used to design primary use clinical systems is

addressed at number 5 at Section 2.1 (Health and care organisations requirements) of this document.

5.2.6 Data quality

The two areas for consideration are the quality of data submitted by providers and the data quality processes at the NCRS offices.

NHS providers:

Each provider is responsible for ensuring the data submitted to the NCRS or submitted through other standard NHS routes is of the highest quality and completeness possible, and accurately represents the service provided.

The NCRS will provide a dynamic feedback process from the Encore cancer registration system to providers. This will allow data quality assurance at a field level – with clinical teams given secure access to the data that their organisation has submitted. Although this will be through a separate web-based interface the NCRS will also provide an application programming interface (API) for system suppliers who wish to integrate the data quality assessments into their data collection systems. The secure access to the data will be controlled by the National Cancer Registration Service and not devolved to system suppliers.

National Cancer Registration Service:

One of the main roles of the NCRS is to ensure data quality and consistency. The eight NCRS regional offices have now moved to a single online processing system (Encore). The working practices will be standardised with continuous performance monitoring and oversight of the entire NCRS through the NCIN and Public Health England.

Specific aspects of data quality are described in Appendix G.

5.2.7 Demographic data

The cancer registration dataset is dynamic and individual tumour records are updated from numerous disparate data sources. Data quality of these sources across the NHS is not sufficiently good to allow accurate mapping of new data to existing items without patient-identifiable data. Even once linked, retaining addresses and names remains important; the address stored by the NCRS is that at the time of diagnosis of the tumour and is essential for cancer cluster analyses possibly many years later, when the patient may have moved. Without

patient name, registries could not support genetic and follow-up enquiries made by clinicians who often only have limited information on the index case and possible relatives.

Nevertheless it is registry practice to use pseudonymised or even anonymous (possibly still disclosive) datasets for analysis where patient identity in not needed. Access to identifiable and potentially disclosive data requires appropriate permissions from the NHS Health Research Authority.

For details on how the cancer registry processes deal with linkage and data discrepancies please see Appendix H.

5.3 Technical architecture

5.3.1 Implementation overview

The COSD modifies and extends the pre COSD requirements on NHS providers to submit monthly cancer data returns to the NCRS by approved NHS secure methods. Providers should therefore have reviewed and revised their previous arrangements to submit monthly returns to their local NCRS office in relation to the timeframe, content and format of those returns in order to conform to this specification.

It is recognised that the data items may be recorded in different electronic systems and there is no requirement to send an integrated record of all data items in one file provided the rules for identifying and linking records are followed to enable the data to be recombined by the NCRS.

The pre COSD cancer data returns to the NCRS covered generic (core) data items only, however, in addition to revising this core dataset, the COSD also includes 12 site specific datasets consisting of data items which have been identified by the NCIN site specific clinical reference groups (SSCRGs) as essential for the analysis of outcomes and services of the relevant tumour sites. These include both clinical and pathology data items. The COSD also defines a subset of data items within the core and breast datasets which should be completed for recurrent/secondary breast cancer. Implementation of the standard has been phased over two years by which time the full dataset should be submitted using XML format for all new primary cancers and recurrent/secondary breast cancers. There was a minimum requirement to submit the core dataset and site specific stage data for all new cancer diagnoses and breast recurrences from the start of the implementation, followed by other site specific clinical items and then site specific pathology items.

Providers are therefore expected to update their data extraction processes as per the defined standard. All new extracts should be developed using the XML schema provided as part of the standard. All extracts from MDT management systems should now be submitted in XML. The original deadline for submissions in XML from pathology systems has now been extended to 1 January 2016. Providers will need to develop an action plan for achieving this and to share this with the NCRS.

It is recognised that some changes to pre-existing working practices may be required, particularly in relation to the electronic capture of site specific clinical data items by clinical Multidisciplinary Team (MDT) members in order to facilitate extraction for COSD.

There are no changes to arrangements for submission of data under other approved standards as a result of implementing this standard.

From 1 April 2015 providers are expected to submit records for all cancer recurrences as specified in the guidance documents. In order to meet varied development timescales this is phased over three months and all providers should be meeting this requirement by 1 July 2015.

Please see implementation guidance documents for further details.

5.3.2 Changes to prior data collection/data flows

About two thirds of the data items to be submitted directly by NHS Providers under COSD are covered by pre-existing data collections. Most of these are collected either monthly or at point of care.

The following summarises the changes to these pre-existing collections and flows:

Previous	New
Cancer registration (monthly submission under national contract)	Replaced by COSD monthly submission to NCRS under national contract. Extended to include site specific clinical and site specific pathology items. Submit data from MDT management systems in XML by January 2015
Pathology reports (monthly	Continue to submit to local NCRS

submission to cancer registries to support registration)	office. **Submit in XML by 1 January 2016**
RCPath core pathology datasets (data items covered by professional standard minimum datasets (MDS)	No change – continue to record (subsets included in COSD) **Submit in XML by 1 January 2016**
Cancer waiting times (available to cancer registries monthly, following local validation and central reporting)	No change
National cancer audits	No change. It is expected that data items shared with the national audits will continue to be dual flowed until the quality and completeness of these items through the COSD standard is considered adequate. At this time the audits will be able to obtain the items through the NCRS service and will no longer need to collect them directly

5.3.3 Actions required

NHS Providers will need to:

- identify items in COSD which were previously collected electronically and those that require changes to systems in order to collect them
- identify items in COSD which may be received by the NCRS offices from other routes (eg image exchange portal)
- identify electronic source for all data items. This may include patient administration systems (PAS), MDT software systems, pathology and imaging systems etc
- identify the items that can be extracted in XML format and submitted by the current NHS compliant methods to the NCRS
- identify data items which are not recorded in a structured format but could be submitted in other formats until appropriate electronic software is available. (This is likely to be pathology/imaging data items which are included in pathology/imaging reports)
- agree the method of submission for all data items with the NCRS and record this in a data transfer and partnership agreement (DTPA)

- develop an action plan to ensure extracts from pathology systems can be submitted in XML format by 1 January 2016

A range of supporting documents is available from the COSD pages on the NCIN website. Additional support is available to providers from the NCRS data liaison team. The standard DTPA is held by the NCRS and will be adapted if necessary for individual Providers.

5.3.4 Data sources

Data may be recorded in a variety of systems such as MDT software and patient administration systems and therefore multiple data extract files may be submitted from a variety of sources.

Additional data sources may not have technical capacity for extraction into XML file format so data may be submitted in other formats where necessary and as agreed with the NCRS.

The following diagram shows the data flows to complete the full COSD dataset as it is expected to be by January 2015. (The grey shaded areas are data items in the reference – other standards and other sources sections of the dataset). The coloured items show expected data sources for NHS provider data extracts. There may be some variation in the data sources between providers as configurations of data collection systems are not uniform across the provider community.

Note that the structured flow of pathology and radiology items is subject to development of appropriate systems within NHS providers.

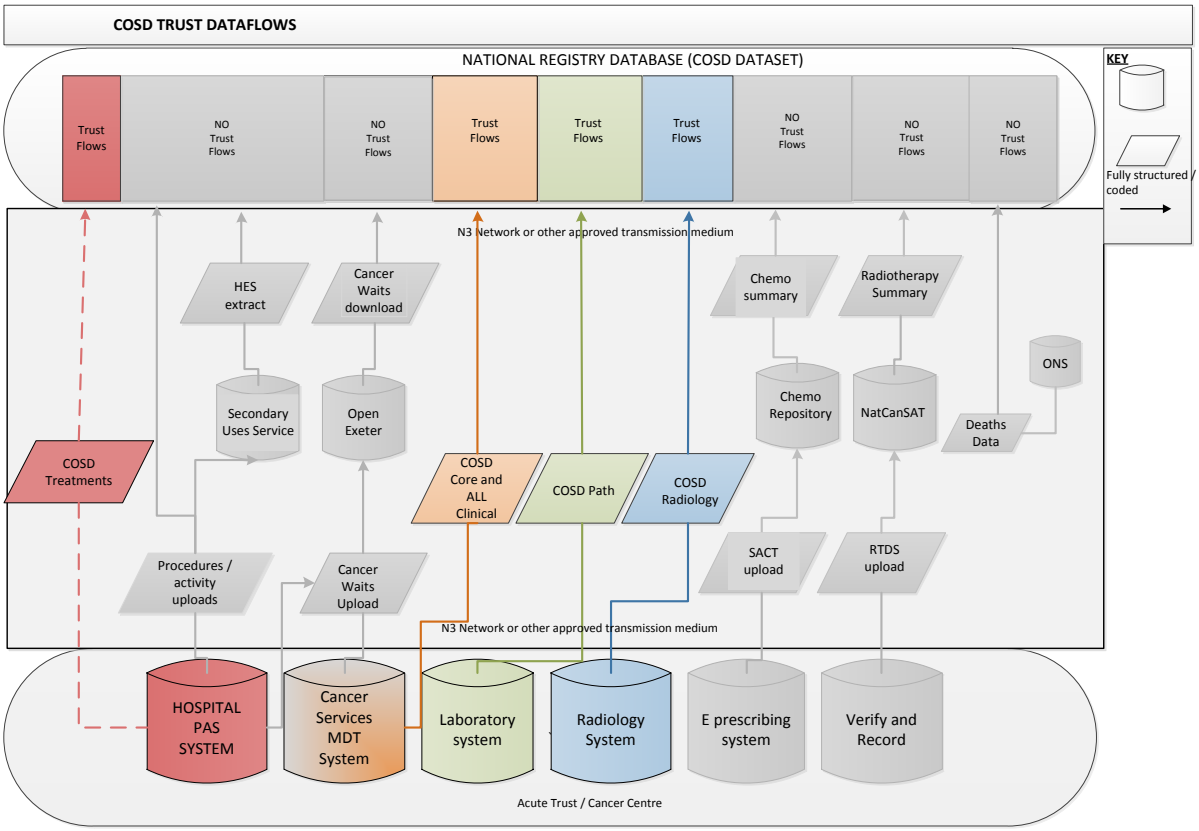


Fig 7: Example of data flows to complete COSD

5.3.5 Submission of data

Providers will submit the data to the NCRS monthly in XML format where available using the current NHS approved standards of submission (see implementation guidance for further details).

5.3.6 XML format for submissions

By January 2015 all the data submissions from MDT management systems should be in the NHS prescribed XML format. It is recognised that this may present a challenge for pathology systems and therefore an extension to 1 January 2016 has been agreed for pathology. Providers are asked to provide an action plan to the NCRS by July 2015 for meeting this requirement.

5.3.7 Phased approach to implementation for dataset v6.0

The original COSD dataset (v1.0) was required to be submitted in full from January 2014, with v1.2 2014 refinements from April 2014. Data extracts from MDT management systems have been required to be submitted in XML from January 2015.

The revised dataset v6.0 is expected to be submitted by July 2015 at the latest with a phased implementation from April 2015. This is to make allowance for the varied timescales of different software suppliers and inhouse developers. During this three month period data can be submitted in accordance with either dataset v1.2 (schema v5-0-4r) or dataset v6.0 (schema v6.0).

Data extracts from pathology systems should be submitted in XML by 1 January 2016 and Providers should submit an action plan to the NCRS meanwhile to describe how they will meet this requirement.

By date	All providers
1 January 2014	**Full COSD dataset (v1.0) submitted**
1 April 2014	**Full COSD dataset (v1.2 Revisions) submitted** **Plans in place to submit data in XML**
1 January 2015	**COSD dataset v1.2 extracts from MDT management systems submitted in XML**
1 April 2015	**First submissions of COSD dataset (v6.0 Revisions)** **First plans sent to NCRS for submission of pathology data in XML**
1 July 2015	**COSD dataset (v6.0 Revisions) submitted** **Any outstanding plans sent to NCRS for submission of pathology data in XML**
1 January 2016	**Data from pathology systems submitted in XML**

5.3.8 Working practices

The implications of the data standard to data providers are as follows;

- NHS providers and system suppliers need to include the new and changed data items in their electronic systems
- these organisations may need to amend their transmission methods to enable the new and changed data items to flow and be centrally collated by the NCRS

- there may be training implications for staff given changes to data item definitions or the implementation of new data items
- provider multidisciplinary teams may need to adjust their previous processes for capturing data in order to include all the data items in the monthly extracts and ensure accuracy of clinical items

5.3.9 Implementation guidance

Implementation guidance continues to be developed to support users, organisations and systems suppliers to implement the standard and updated versions of the documentation are available on the COSD pages of the NCIN website. The previous cancer dataset manuals have been replaced by the COSD User Guide.

6. Supporting information

6.1 Contact details

Information including the COSD dataset and COSD user guide is available on the NCIN website at:

www.ncin.org.uk/collecting_and_using_data/default.aspx

Queries regarding this document should be addressed to:

cosd@ncin.org.uk

Queries regarding submissions should be discussed with the NCRS data liaison team.

Appendix A – mandatory registerable conditions (from the UKACR Library of Recommendations)

ICD 10	Description of neoplasm
C00-C97	All malignant neoplasms
D00-D09	All carcinoma in-situ
D32-D33 D35.2 & D35.3 D35.4	Benign neoplasms of brain & other parts of nervous system Benign neoplasms of pituitary gland & craniopharyngeal duct Benign neoplasms of pineal gland
D37-D48 (excluding D47.2)	All neoplasms of uncertain behaviour Neoplasms of unspecified nature of bladder Neoplasm of unspecified nature of brain Neoplasm of unspecified nature of other parts of nervous system and pituitary gland only (Excluding D47.2 Monoclonal gammopathy of undetermined significance (MGUS))
E85.9	Primary Amyloidosis (Subject to confirmation by the UKIACR)[2]

[2] Although Primary amyloidosis (E85.9) is listed as an E ICD code in the World Health Organization (WHO) disease classification, amongst clinicians it is widely acknowledged and subsequently treated as a cancer, receiving chemotherapy in cases. While we await the WHO disease classification being updated to reflect this, we have extended the scope of the COSD to include this.

(Please see COSD user guide for full list of mandatory registerable conditions (from the UKACR Library of Recommendations)).

Appendix B – NCRS/CRUK patient information leaflet

Registry information has shown that 1 in 2 people will now survive cancer for at least 10 years.

Achievements made possible by cancer registration information:

- research showing that there are at least 10 different types of breast cancer, which means treatments can be made more specific for each type
- monitoring whether cancers are becoming more or less common – for example spotting the rapid increase in skin cancer cases has led to prevention campaigns to promote staying safe in the sun and avoiding sun beds
- improvement of the breast cancer screening programme, and the decision to introduce flexible sigmoidoscopy (a technique for examining the bowel) as a method of screening for bowel cancer
- research around when and where patients are diagnosed with cancer, which showed that almost a quarter are diagnosed in an emergency. This has reinforced the importance of finding ways to get more patients diagnosed early

The more information we have in the registry, the easier it is to improve diagnosis and treatment.

What if I don't want my details on the cancer registry?

The benefits of the data collected by the cancer registry have been considerable and we are grateful that nearly everyone with cancer is prepared to share their data with the cancer registry. However, you can ask us to remove all of your details from the cancer registry at any time. These requests won't affect your treatment or care. If you wish to make such a request, you should email **optout@ncr.nhs.uk** or write to

Director
National Cancer Registration Service
Public Health England
Wellington House
London SE1 8UG

If you have any questions about cancer registration, you can get more information by:

- asking your doctor
- visiting the Cancer Research UK website at **www.cruk.org/cancer-registration** or the cancer registration website at **www.ncr.nhs.uk/patientinfo** where you will find a longer booklet
- and for any questions on cancer, speak to one of Cancer Research UK's nurses on freephone **0808 800 4040** (9am–5pm, Monday to Friday)

Cancer registration is crucial for progress against cancer, and is supported by all the main UK cancer charities and cancer patient groups.

Against Breast Cancer
Bowel & Cancer Research
Bowel Cancer UK
Brain Tumour Research
Brain Tumour Research Campaign
Brainstrust
Breast Cancer Campaign
British Lung Foundation
Core – the Digestive Disorders Foundation
Cancer52
Cancer Research UK
GIST Support UK
It's in the Bag
James Whale Fund for Kidney Cancer
Jo's Cervical Cancer Trust
Skcin – The Karen Clifford Skin Cancer Charity
Leukaemia & Lymphoma Research
Lymphoma Association
Macmillan Cancer Support
Marie Curie Cancer Care
Melanoma Focus
My Name is NOT Cancer
Myeloma UK
Pancreatic Cancer Action
Rarer Cancers Foundation
Sarcoma UK
Shine Cancer Support
Skin Cancer Research Fund
Target Ovarian Cancer
Teenage Cancer Trust
The Pelican Cancer Foundation
The Pink Ribbon Foundation
WMUK

PHE publications gateway number: 2014449

Public Health England

Protecting and improving the nation's health

Cancer Registration – what it is, the benefits of being on the register, and your options

NHS CANCER RESEARCH UK

Appendix C – UKACR policy on requests from patients to delete their information

Patient opt out request form

This form is for use by patients to request that their personal information be excluded from processing onto the National Cancer Registration Service's cancer registration database.

The personal information collected on this form is needed so that we can process your request correctly. It will only be used in connection with carrying out this request.

To be completed by the patient. Please complete as fully as possible.

<table>
<tr><td>My details</td><td></td></tr>
<tr><td>My name</td><td></td></tr>
<tr><td rowspan="5">My address</td><td>NHS number</td></tr>
<tr><td>Date of birth</td></tr>
<tr><td>Sex</td></tr>
<tr><td>Telephone number</td></tr>
<tr><td>Place last treated</td></tr>
<tr><td>Post code</td><td>Date last treated</td></tr>
</table>

My request

I wish the cancer registration system in England to stop adding information about me to the cancer registration database and either

☐ Delete everything except for 'My details'

Remove, as far as possible all clinical information relating to me, but **retain** my NHS number and the information I have provided in the 'My details' section above in the 'watch list', so that any further information

received about me will not be processed by the national cancer registration system.

or:

☐ Delete everything
Remove, as far as possible all clinical and personal information relating to me **including** my NHS number and the information I have provided in the 'My details' section above. I understand the registry will not keep any record of my details, so, will not know that any information received about me in the future should not be processed.

We will send you a copy of the leaflet, 'Cancer Registration – what it is, the benefits of being on the register, and your options', in the hope that you may change your mind about opting out. If you do change your mind then please contact us as soon as you can.

Signed _____ Date _____

Please return to: Dr Jem Rashbass
Director,
National Cancer Registration Service
Public Health England
Wellington House
London
SE1 8UG

Patient advice

The law permits the National Cancer Registration Service (NCRS) to collect information on all cancer patients in England.

The law also gives you, the patient the right to opt out of cancer registration. The first stage of opt out is usually a discussion with a clinician. The clinician will ensure that you are fully aware of the value of your information to research and for improving cancer treatments.

If you still want to opt out and have your details removed from the cancer registration data, you must apply in writing preferably by using the 'patient opt-out request' form in Appendix 1.

The NCRS will then add your NHS number to an ‘exemption list’. By adding your details to this list, the NCRS can ensure that it will not collect any incoming information about you.

The NCRS will search all cancer registration files and records and as far as is practicable delete any existing information relating to you that it may already have.

The NCRS will also check whether it has sent any identifiable information to other permitted organisations such as the Office for National Statistics, and if so as far as is practicable contact that organisation and instruct them to delete the information.

The NCRS will complete its actions within 20 days of receiving the written request and will confirm this in writing to you.

You may also request removal from the exemption list and the NCRS will act on this request, however if your details are removed from the exemption list, the NCRS will not be able to guarantee that your data is not added in future.

Appendix D – UKACR policy on data retention and disposal

Implementation: 1 September 2003.

The Data Protection Act 1998 places statutory restrictions on the use of personal information, which should not be kept longer than is necessary for the purpose.

Background

One of the prime purposes of the cancer registry is to produce information on cancer over long time periods. Historic information is essential to this role. Whilst many of the uses of this historic data do not require access to identifiable data, there are particular instances where identifiable data must be used.

This document describes:

- what data are retained, and for how long
- reasons why the data are retained, and their uses
- how registries control access to retained data

What Data are Retained?

Currently:

- all identifiable data on cancer patients held by cancer registries are held indefinitely
- there is no disposal of identifiable data for cancer patients

Reasons for Retention of Data

In many organisations the ‘usefulness’ of data to the organisation diminishes over time. In such instances the organisation reaches a time-point when it is no longer appropriate to retain those data. This is not true for data held by cancer registries. For significant reasons, the older cancer registration data remain important and retaining identifiable data is of increased importance.

One of the main purposes of cancer registration is to monitor cancer survival over time and this requires the ability to reliably link data from multiple sources (including ultimately death certificates) over potentially long periods of time. NHS number is not yet routinely present on death certificates, so that linkage is only possible using other identifiers. People dying now of or with cancers of good prognosis diagnosed decades ago, may not necessarily have NHS numbers on their registration records. Again, the only way to link these records is by using other identifiers.

An important use of cancer registry data of direct clinical care relevance is the tracing of family members with cancer related to an index case (proband), who may be a cancer patient or may be cancer-free but concerned about family risk. These requests originate from Clinical Genetics Services. This important service was recognised and provided for in paragraph (1) (e) of Regulation 2 of the Statutory Instrument 2002 No 1438. It is also explicit in both the UKACR Confidentiality Guidelines (www.ukacr.org.uk) and the data release policy [Item 1)c)], which was approved at the PIAG meeting of March 2003.

Not only have these requests increased rapidly in recent years (by around 15 times since 1996), indicative of their usefulness to clinical geneticists, but requests to any registry can originate anywhere in the UK and often relate to family members who died or were diagnosed with cancer early in the 20th century. To enable accurate identification or non-identification it is imperative to have as much identifying details on all cancer patients as possible.

In accordance with the UKACR guidelines, registries release this information only if the individuals concerned have given their consent (if alive), and then it is released only to the clinical geneticist in person. In many cases medical records related to these individuals will have been destroyed and cancer registries are the only source of information. The usual requirement is for the cancer registries to confirm the age at diagnosis and the site of cancer in the relative. This is not possible without the use of identifiable data.

The increased number of these requests nationally reinforces that this is a reliable, effective and controlled mechanism for the identification of cases for use within genetic counselling. It is essential that cancer registries maintain a positive approach to supporting genetic clinics in assessment of familial risk. Without the ability to search back through time using identifiable data, this type of service would not be available.

As research evidence grows, it may become apparent that certain patients treated for a particular type of cancer can be exposed to a much higher degree

of risk of developing secondary cancers. Identifiable cancer registry data are recognised as the prime mechanism in the facilitation of a national exercise of selective patient notification. Currently the Chief Medical Officer is asking the Cancer Registration Service to help to compile information for a notification exercise on a specific group of cancer patients at an increased risk of developing second cancers. All such programmes would have to comply with the strict Policy on Data Release.

Appendix E – UKACR policy on data disclosure

(The policy shown below will shortly be available on the UKIACR website (http://www.ukiacr.org/) which has replaced UKACR website)

Home » UKACR Disclosure Policy

UKACR Disclosure Policy

UKACR Policy on disclosure of identifiable data by cancer registries – guidance on implementation within England and Wales

Background

Regulation 2 of the Statutory Instrument (SI) on confidentiality – No. 1438, The Health Service (Control of Patient Information) Regulations 2002 – permits cancer registries to receive patient identifiable data [note 1] without the need for informed consent and it permits registries to process said data for the medical purposes stipulated in regulation 2. The regulation was made under Section 60 of the Health and Social Care Act 2001 and continues to have effect under Section 251 of the NHS Act 2006. The approval has been subject to annual review by the Patient Information Advisory Group (PIAG). The functions of PIAG have now been taken over by the Ethics and Confidentiality Committee (ECC) of the National Information Governance Board (NIGB).

When dealing with requests for patient identifiable data registries must assess each request carefully and on merit, to ascertain whether or not patient identifiable data are really necessary. If not, anonymised data must be supplied, following the relevant UKACR guidelines. This is an important principle that registries must apply, even if patient identifiable data have been provided for similar requests in the past.

The UKACR guidance is informed by the Disclosure Review for Health Statistics (referred to as the Health Review) developed by the Office for National Statistics and approved by ministers as policy in England. The health review provides detailed guidance on how to decide whether or not data are identifiable, and is the standard reference for publishing health data.

Summary

Cancer registries in England and Wales can release patient identifiable data legally only to those organisations specified in items 1) a), 1) b), 1) c), 1)d) and 1) e. All other organisations or individuals need approval from the Ethics and Confidentiality Committee of the National Information Governance Board, unless they have informed consent from patients. This policy must be implemented by all organisations listed in Appendix 1. This policy will be subject to annual review.
All requests for patient identifiable data must be made using the UKACR request form for patient identifiable or potentially identifiable data or the registry's host organisation's standard request form for identifiable data [note 2].

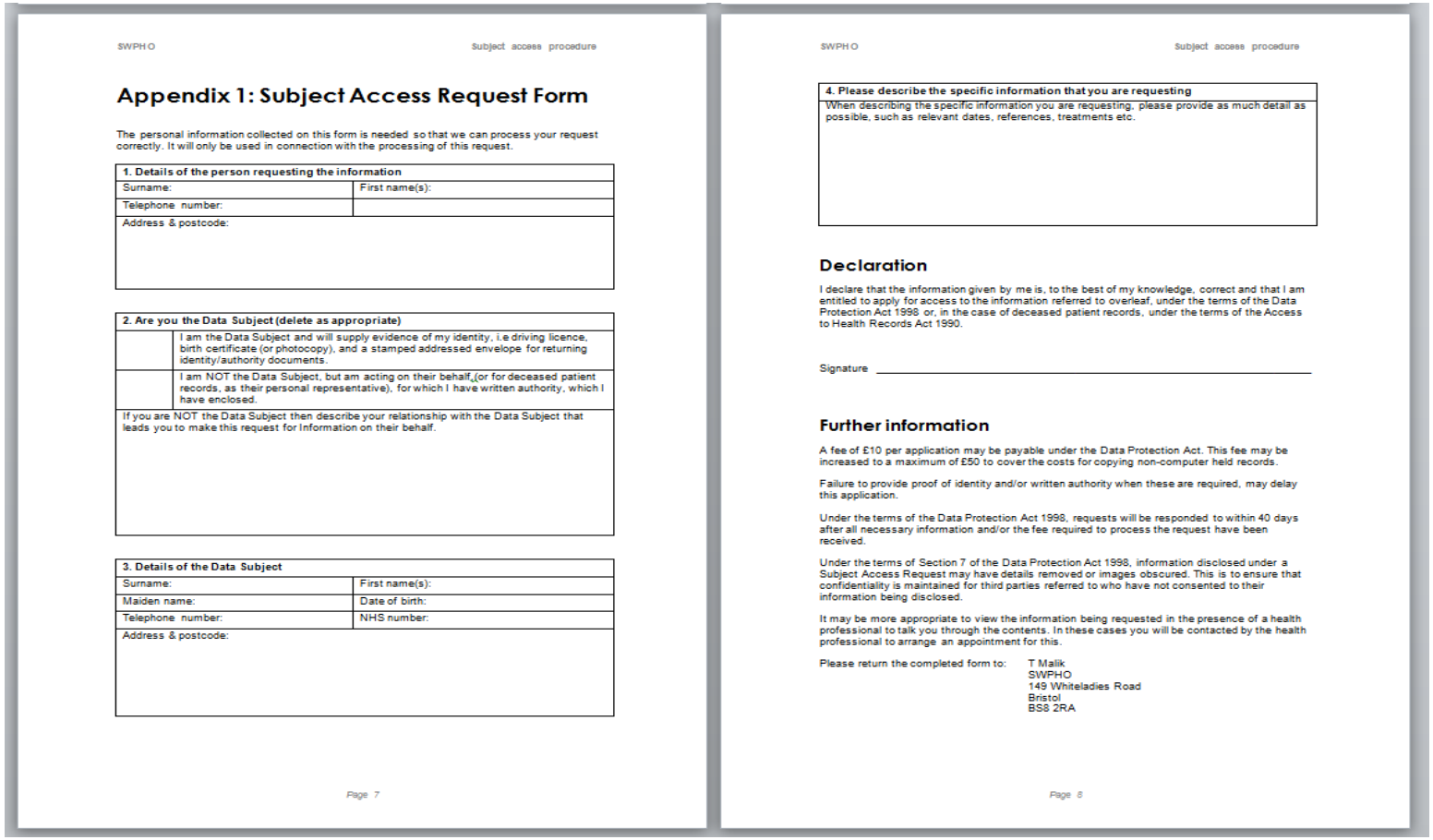
Chris Carrigan, National Cancer Registration Co-ordinator, England
Monica Roche, Co Chair, United Kingdom Association of Cancer Registries

[note 1] Defined in the Health and Social Care Act 2001 as "For the purposes of this section, patient information is "confidential patient information" where:
(a) the identity of the individual in question is ascertainable:
(i) from that information, or
(ii) from that information and other information which is in the possession of, or is likely to come into the possession of, the person processing that information, and
(b) that information was obtained or generated by a person who, in the circumstances, owed an obligation of confidence to that individual.
Data will be regarded as identifiable if it includes any of the following data items: name, address, postcode, date of birth, date of death, NHS number, hospital number.
[note 2] Registry specific versions of the UKACR request form for patient identifiable or potentially identifiable data are available at: (http://www.ukacr.org/confidentiality).

Reference documents

- Health and Social Care Act 2001
  (www.legislation.gov.uk/ukpga/2001/15/contents

Appendix F – a sample subject access request policy from the SWPHO*



SWPHO Subject access procedure

Appendix 1: Subject Access Request Form

The personal information collected on this form is needed so that we can process your request correctly. It will only be used in connection with the processing of this request.

1. Details of the person requesting the information	
Surname:	First name(s):
Telephone number:	
Address & postcode:	

2. Are you the Data Subject (delete as appropriate)	
	I am the Data Subject and will supply evidence of my identity, i.e driving licence, birth certificate (or photocopy), and a stamped addressed envelope for returning identity/authority documents.
	I am NOT the Data Subject, but am acting on their behalf,(or for deceased patient records, as their personal representative), for which I have written authority, which I have enclosed.
If you are NOT the Data Subject then describe your relationship with the Data Subject that leads you to make this request for Information on their behalf.	

3. Details of the Data Subject	
Surname:	First name(s):
Maiden name:	Date of birth:
Telephone number:	NHS number:
Address & postcode:	

Page 7

SWPHO Subject access procedure

4. Please describe the specific information that you are requesting
When describing the specific information you are requesting, please provide as much detail as possible, such as relevant dates, references, treatments etc.

Declaration

I declare that the information given by me is, to the best of my knowledge, correct and that I am entitled to apply for access to the information referred to overleaf, under the terms of the Data Protection Act 1998 or, in the case of deceased patient records, under the terms of the Access to Health Records Act 1990.

Signature _____

Further information

A fee of £10 per application may be payable under the Data Protection Act. This fee may be increased to a maximum of £50 to cover the costs for copying non-computer held records.

Failure to provide proof of identity and/or written authority when these are required, may delay this application.

Under the terms of the Data Protection Act 1998, requests will be responded to within 40 days after all necessary information and/or the fee required to process the request have been received.

Under the terms of Section 7 of the Data Protection Act 1998, information disclosed under a Subject Access Request may have details removed or images obscured. This is to ensure that confidentiality is maintained for third parties referred to who have not consented to their information being disclosed.

It may be more appropriate to view the information being requested in the presence of a health professional to talk you through the contents. In these cases you will be contacted by the health professional to arrange an appointment for this.

Please return the completed form to: T Malik
SWPHO
149 Whiteladies Road
Bristol
BS8 2RA

Page 8

South West Public Health Observatory is part of Public Health England from 1st April 2013

Appendix G – NCRS data quality controls

Automated quality control.

The data files submitted through the NCRS data clearing house are subject to a wide range of validation rules to ensure that the data files and data within fields is consistent, as follows:

- **batch tracing of all cases** - all patient-identifiable electronic records are sent to the demographics batch service (the replacement for NSTS) for tracing against the NHS spine discrepancies are identified, investigated and whenever possible reconciled
- it is not anticipated that this will place any additional load on either the personal demographic service (PDS) or batch tracing facilities, however, it is acknowledged that use of this service is within purchased volumetric limits as the Health and Social Care Information Centre does not support the creation of new databases using batch tracing - in this case the system should consider full PDS compliance; the impact will be monitored by the NCRS as the project progresses with any significant increase being brought to the attention of the COSD Project Board

- **use of multiple data sources** - the quality of cancer registration relies upon the use of multiple independent data sources to ensure high ascertainment and cross validation - the Encore system automates much of the data linkage between the disparate sources – highlighting inconsistencies that can be further investigated

- **cancer registration staff** - the National Cancer Registration Service employs tumour registration staff at the local registries; the registration staff have considerable expertise in cancer coding and classification and spend much of their time quality assuring the electronic data sources and cases recorded at the registry - in some cases cancer registration officers work in a local provider organisation, but all registries maintain very close contacts with the clinical teams

- **data feedback to clinical teams** - rapid feedback to the provider clinical teams, usually through the MDT or Cancer Programme Board provides an important process of data validation - the NCRS uses secure web-based systems to deliver reports at a field-level on the completeness of individual data items

- **data quality audit** - the UKACR has developed a large number of performance metrics covering the process of data collection by registries; these performance metrics have been integrated into the new Encore system and will, where appropriate form the basis of daily updates on the data quality and completeness of records held in the NCRS

Appendix H – data linkage and data discrepancies

Linkage

Linkage is a complex issue, which has become far simpler in recent years with the rollout in use of the NHS number. Registries use different linkage methods according to the type of data which is available. In essence, the more data that is available, the more confident that linkage is correct.

In fact linkage comprises two parts; blocking and weighting. Blocking takes an incoming record and uses a range of search criteria, determined by the incoming records content, to identify a series of possible matches in the database. Where the NHS number is available, that is used, but other blocking is usually also applied. In a manual context, these blocks tend to be sequential, but in an automated setting they tend to run consecutively, with all potential matches passing to the second stage, weighting.

Weighting can be simple. Deterministic weighting is used for NHS number matching, but this is always augmented with at least one other identifier. Probabilistic techniques use a wider set of data matches, and are usually used when the NHS number is not available on either the source record or the blocked record. It looks for the ‘commonness’ of the data value in the overall database, and then uses that to weight up or down based on a series of random control matches. Probabilistic weighting is a well defined science, with robust methodologies, however it is used far less than in past years.

Data discrepancies

The fundamental principle of cancer registration is that it relies on multiple sources of data. When dealing with multiple sources, many of which may contain a common item, there is likelihood that two sources will give different values for a particular item of data.

The technical design of the registration schema is such that multiple sources and multiple data values are held against the summarised registration record. Registration clerks are trained to identify and deal with data discrepancy. This usually starts with some basic data checking with the source data supplier, but where conflicting data exists there are clear rules by which registry staff undertake this. At no stage is any source data overwritten or lost, and regular checks are included in the registration practice to examine random sets of records as part of the standard QA built into registration practice over many years.

Appendix I – approval under Health Service (Control of Patient Information) Regulations 2002

Application number	0001	
Reference	PIAG 03(a)/2001	
Other refs		
Application title	National cancer registry databases	
Application summary	Application submission by United Kingdom Association of Cancer Registries (UKACR) represented several different registries with a common purpose, which is to use patient information on cancer registry databases.	
Applicant organisation name	United Kingdom Association of Cancer Registries (UKACR)	
Contact name	Roger Hartley, head of cancer registration, National Cancer Registration Service (North West)	
Address	Public Health England, The Palatine Centre	
	63-65 Palatine Road, Withington	
	Manchester	
Postcode	M20 3LJ	
Telephone	0161 446 3584	
Email	roger.hartley@phe.gov.uk	
Medical purposes	Y	preventative medicine
	□	medical diagnosis

	□	medical research, approved by a research ethics committee
	□	the provision of care and treatment
	□	the management of health and social care services
	□	informing individuals about their physical or mental health or condition, the diagnosis of their condition or their care and treatment
Cohort/population	UK-wide patients diagnosed with cancer	
Description of confidential patient information used		
S251 class(es)	Y	Specific support
	□	Class I - making the person less readily identifiable
	□	Class II - present or past geographical locations of patients
	□	Class III - to identify and contact patients to obtain consent
	□	Class IV - linking multiple sources; validating quality and completeness; avoiding error
	□	Class V - audit, monitoring, and analysis of healthcare provision
	□	Class VI - granting of access to data for purposes I-V
Sponsor		

Status	Approved
Outcome date	
Next review date	25/11/2015
Notes	Following consideration and referral from the June and August 2014 meetings, a recommendation of continuing support was deferred while Public Health England seek advice from the Information Commissioner's Office on issues of fair processing. Progress on actions arising from these discussions in conjunction with a number of outstanding aspects are to be considered on 06 November 2014. Support continues while this consideration is ongoing. This application was approved by Parliament in Statutory Instrument 2002 No. 1438: The Health Service (Control of Patient Information) Regulations 2002